AD 680437

THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

A NOTE ON THE EVALUATION OF A MULTIVARIATE NORMAL INTEGRAL BY THE METHOD OF DAS

by

J. T. WEBSTER

D D C
RECEIVED
JAN 14 1969
B

Technical Report No. 15
Department of Statistics THEMIS Contract

This document has been approved for public release and sale; its distribution is unlimited

Distribution of this document is unlimited.

Department of Statistics
Southern Methodist University
Dallas, Texas 75222

8

THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

A NOTE ON THE EVALUATION OF A MULTIVARIATE NORMAL INTEGRAL BY THE METHOD OF DAS

by

J. T. Webster

Technical Report No. 15
Department of Statistics THEMIS Contract

December 3, 1968

Research sponsored by the Office of Naval Research
Contract N00014-68-A-0515
Project NR 042-260

Reproduction in whole or in part is permitted
for any purpose of the United States Government.

Distribution of this document is unlimited.

DEPARTMENT OF STATISTICS
Southern Methodist University

A note on the evaluation of a multivariate normal integral by the method of Das

By J. T. Webster
Southern Methodist University

1. Introduction

Das (1956) presents a method of evaluating the integral

$$I = \int_{a_1}^{\infty} \cdots \int_{a_n}^{\infty} f(x_1 , x_2 , \cdots , x_n)\, dx_1 dx_2 \cdots dx_n$$

(where $f(x_1 , x_2 , \cdots , x_n)$ is the joint multivariate normal density function with zero means and nonsingular variance-covariance matrix Σ) through the combining of $n + k$ independent normal variables with zero means and unit variances. Later Marsaglia (1963) shows that this is a special case of a convolution formula. The complexity of implementing the solution is highly dependent upon the size of k and Marsaglia (1963) notes that k equal to n minus the multiplicity of the smallest latent root of Σ can always be achieved. This note investigates properties of Σ that will allow smaller values of k .

2. The equivalent expression

As a method for evaluating I , Das (1956) considers two row rectors $\underline{y}' = (y_1 , y_2 , \cdots , y_n)$ and $\underline{z}' = (z_1 , z_2 , \cdots , z_k)$ all of whose elements are normally and independently distributed with zero means and unit variances. The problem is to choose a positive constant c and an $n \times k$ real matrix B such that

$$\Sigma = c^2 I_n + BB' \quad ; \qquad 1)$$

for then $\underline{x}' = (x_1 , x_2 , \cdots , x_n)$ can be expressed as

$$x = c\underline{y} - B\underline{z} .$$

I can now be expressed as

$$\begin{aligned} I &= \Pr(x_1 \geq a_1 , x_2 \geq a_2 , \cdots , x_n \geq a_n) \\ &= \Pr\left[y_1 \geq \left(a_1 + \sum_{j=1}^{k} b_{1j}z_j\right)/c , y_2 \geq \left(a_2 + \sum_{j=1}^{k} b_{2j}z_j\right)/c , \cdots \right. \\ &\qquad \left. y_n \geq \left(a_n + \sum_{j=1}^{k} b_{nj}z_j\right)/c\right] \\ &= (2\Pi)^{-k/2}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \prod_{i=1}^{n} P\left[\left(a_i + \sum_{j=1}^{k} b_{ij}z_j\right)/c\right] \exp\left(- \underline{z}'\underline{z}/2\right) \prod_{j=1}^{k} dz_j . \qquad 2) \end{aligned}$$

where $P[a] = \Pr(y_1 \geq a)$. It is evident that a small k is advantageous.

The slight change in 1)

$$\Sigma = C^2 + BB' , \qquad 3)$$

where C is a diagonal matrix with positive diagonal elements c_i , results in

$$\underline{x} = C\underline{y} - B\underline{z}$$

and 2) of the form

$$I = (2\Pi)^{-k/2}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \prod_{i=1}^{n} P\left[\left(a_i + \sum_{j=1}^{k} b_{ij}z_j\right)/c_i\right]\exp\left(- \underline{z}'\underline{z}/2\right) \prod_{j=1}^{k} dz_j . \qquad 4)$$

3. Latent vector properties

If Σ can be expressed in the form 3) and $\underline{w}_\ell$ is a latent vector

of Σ then

$$\begin{aligned}\lambda_\ell \underline{w}_\ell &= \Sigma\, \underline{w}_\ell \\ &= C^2 \underline{w}_\ell + BB'\underline{w} \\ &= C^2 \underline{w}_\ell + \sum_{j=1}^{k} \underline{b}_j \underline{b}_j' \underline{w}_\ell \end{aligned} \qquad 5)$$

where $\underline{b}_j$ is the j^{th} column of B . From expression 5)

$$\lambda_\ell w_{i\ell} = c_i^2 w_{i\ell} + \sum_{j=1}^{k} \theta_{j\ell} b_{ij} \qquad 6)$$

where $w_{i\ell}$ is the i^{th} element of $\underline{w}_\ell$ and $\theta_{j\ell} = \underline{b}_j' \underline{w}_\ell$. If the c_i^2 are not all equal the n equations of 6) are parametric equations for a k dimension variety in n space of order two, V_k^2 ; a brief discussion of this may be found in Kendall (1956), page 5. That is given a vector $\underline{w}_\ell$, algebraic operations to eliminate λ_ℓ and the $\theta_{j\ell}$ will result in n - k second degree homogeneous equations in the $w_{i\ell}$. If the c_i^2 are all equal say to c then 6) can be written as

$$w_{i\ell} = \sum_{j=1}^{k} \theta^*_{j\ell} b_{ij} \quad . \qquad 7)$$

where $\theta^*_{j\ell} = \underline{b}_j' \underline{w}_\ell / (\lambda_\ell - c^2)$, which are the parametric equations of a k dimension flat, S_k .

Conversely if the latent vectors of Σ all fall in a V_k^2 of the form of 6) with

$$\theta_{j\ell} \underline{w}_\ell' \underline{b}_j' \geq 0 \text{ and } \lambda_\ell \geq c_i^2 > 0 \qquad 8)$$

then Σ can be written in the form 3) . If the latent vectors of Σ all

fall in a S_k with

$$\theta^*_{j\ell}\ \underline{w}'_\ell \underline{b}_j \geq 0 \qquad 9)$$

then it can be written in the form 1) .

Thus the smallest value of k for evaluating I through form 4) is determined by the smaller of the minimum dimension of the V_k^2 's of the form 6) satisfying conditions 8) and the minimum dimension of the S_k satisfying the condition 9) that contain the latent vectors of Σ .

Alternative conditions can be given for the special case of $k = 1$. 6) becomes

$$\lambda_\ell w_{i\ell} = c_i^2 w_{i\ell} + \theta_\ell b_i$$

or

$$\frac{1}{w_{i\ell}} = -\frac{1}{\theta_\ell}\frac{c_i^2}{b_i} + \frac{\lambda_\ell}{\theta_\ell}\frac{1}{b_i} \qquad 10)$$

which is the parametric expression of a plane, S_2 , passing through the origin. Thus if the $\underline{s}'_\ell = (1/w_{1\ell}\ ,\ 1/w_{2\ell}\ ,\ \cdots\ ,\ 1/w_{n\ell})$ all fall in a plane with parametric conditions 10) satisfying

$$(\lambda_\ell - c_i^2)\,\theta_\ell\,\underline{w}'_\ell \underline{b} > 0$$

only one auxiliary variable is necessary. For the case of $c_i^2 = c^2$ for all i , n - 1 of the λ_ℓ equal c^2 and their corresponding θ_ℓ are zero. That is the minimum latent root of Σ has multiplicity k - 1 .

4. Summary

In general the minimum number of auxiliary variables necessary to evaluate a multivariate normal integral through the method of Das (1956)

can be determined by the dimensions of the second order varieties with parametric form 6) and the dimensions of the flats that contain the latent vectors of the variance-covariance matrix.

This problem could also be looked at as determining the diagonal matrix D such that $D \Sigma D$ has maximum multiplicity of its smallest latent root. Then Marsaglia's (1963) solution suffices. It is interesting to note that the difficulty of the solution, the dimension of the integration in I , is affected by change of scales of the normal variables.

REFERENCES

Das, S. C. (1956). The numerical evaluation of a class of integrals, II. Proc. Camb. Phil. Soc. 52, 442-448.

Kendall, M. G. (1961). A Course in the Geometry of n Dimensions. London: Charles Griffin and Co.

Marsaglia, G. (1963). Expressing the normal distribution with covariance matrix $A+B$ in terms of one with covariance matrix A . Biometrika, 50, 535-538.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) SOUTHERN METHODIST UNIVERSITY	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
	2b. GROUP UNCLASSIFIED

3. REPORT TITLE
A Note on the Evaluation of a Multivariate Normal Integral by the Method of Das

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report

5. AUTHOR(S) (First name, middle initial, last name)
J. T. Webster

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
December 3, 1968	5	3

8a. CONTRACT OR GRANT NO. N00014-68-A-0515	9a. ORIGINATOR'S REPORT NUMBER(S) 15
b. PROJECT NO. NR 042-260	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT
Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research

13. ABSTRACT

Das (1956) presents a method of evaluating the integral

$$I = \int_{a_1}^{\infty} \cdots \int_{a_n}^{\infty} f(x_1, x_2, \cancel{\ldots}, x_n)\, dx_1 dx_2 \ldots dx_n$$

(where $f(x_1, x_2, \ldots, x_n)$ is the joint multivariate normal density function with zero means and nonsingular variance-covariance matrix Σ) through the combining of $n + k$ independent normal variables with zero means and unit variances. Later Marsaglia (1963) shows that this is a special case of a convolution formula. The complexity of implementing the solution is highly dependent upon the size of k and Marsaglia (1963) notes that k equal to n minus the multiplicity of the smallest latent root of Σ can always be achieved. This note investigates properties of Σ that will allow smaller values of k.

References

Das, S. C. (1956). The numerical evaluation of a class of integrals, II. Proc. Camb. Phil. Soc. 52, 442-448.

Marsaglia, G. (1963). Expressing the normal distribution with covariance matrix A+B in terms of one with covariance matrix A. Biometrika, 50, 535-538.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification